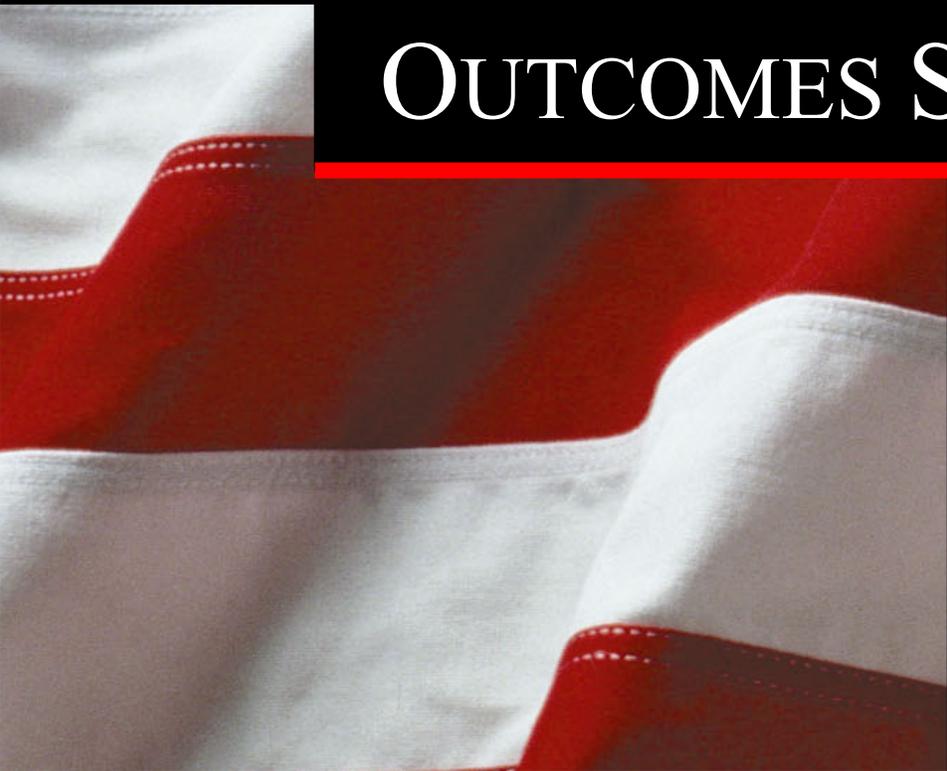




**ELECTRONIC DATA
USER'S GUIDE**

Cohort IV Baseline
2001

MEDICARE HEALTH



OUTCOMES SURVEY

**CENTERS
FOR MEDICARE
& MEDICAID
SERVICES**

**HEALTH
SERVICES
ADVISORY
GROUP**

Table of Contents

PREFACE	1
INTRODUCTION	2
BACKGROUND	
MEDICARE HEALTH OUTCOMES SURVEY INSTRUMENT	
METHODOLOGY	6
SAMPLING METHODOLOGY	
DISTRIBUTION OF THE SAMPLE	
SCORING SF-36 [®] PHYSICAL AND MENTAL HEALTH SUMMARY MEASURES	
CASE MIX ADJUSTMENT	
BASELINE ELECTRONIC DATA FILE CHARACTERISTICS	10
PLAN LEVEL VARIABLES	
BENEFICIARY LEVEL VARIABLES	
SURVEY LEVEL VARIABLES	
PLAN, VENDOR, AND SURVEY STATISTICS	
ANALYTIC VARIABLES	
SF-36 [®] SUMMARY MEASURES	
PACE DISCUSSION	13
APPENDIX A	A1
Quality Assurance of the Data	
DATA RECEIPT	
DATA CLEANING AND EDITING	
APPENDIX B	B1
Baseline Electronic Data File Specifications	

Preface

The Centers for Medicare & Medicaid Services (CMS) is committed to monitoring the quality of care provided by Medicare + Choice Organizations (M+COs). The Medicare Health Outcomes Survey (HOS) is the first health outcomes measure for the Medicare population in managed care settings. The HOS design is based on a randomly selected sample of individuals from each participating M+CO, and measures physical and mental health over a two-year period.

This *Cohort IV Baseline Data User's Guide* is designed to assist individuals with the use of the beneficiary level *Cohort IV Baseline* Electronic Data. The data user's guide includes a general overview of the survey methodology, data cleaning, and scoring processes required for the creation of the HOS *Cohort IV Baseline* data set. These data are intended to support M+CO and Quality Improvement Organization (QIO) quality improvement activities.

Introduction

BACKGROUND

The CMS, in collaboration with the National Committee for Quality Assurance (NCQA), launched the first Medicare managed care outcomes measure in the Health Plan Employer Data and Information Set (HEDIS[®]) in 1998.¹ The measure includes the most recent advances in summarizing physical and mental health outcomes results and appropriate risk adjustment techniques. This measure was initially titled Health of Seniors, and was renamed the Medicare Health Outcomes Survey during the first year of implementation. This name change was intended to reflect the inclusion of Medicare recipients in the sampling methodology who are disabled and under age 65.

The HOS measure was developed under the guidance of a Technical Expert Panel (TEP) comprised of individuals with specific expertise in the health care industry and outcomes measurement. The TEP continues to oversee and develop the science of the HOS measure. The CMS has contracted with NCQA to support the standardized administration of the HOS survey, including selecting, training, and certifying independent survey vendors that the plans contract with to administer the survey.

The HOS measure is an assessment of a health plan's ability to maintain or improve the physical and mental health functioning of its people with Medicare over a two-year period of time. The functional status of the elderly is known to decline over such a period.² The differences between the baseline and the two-year follow up physical and mental health scores are aggregated at the plan level, yielding HOS plan level Performance Measurement results. The *Cohort I* Performance Measurement results were released in 2001. The Performance Measurement results for *Cohort IV Baseline* (scheduled for release in 2004) will incorporate data from the 2003 *Cohort IV Follow Up* survey.

The HOS *Cohort IV Baseline* Report distributed in June 2002 is part of a larger effort by CMS to improve the health care industry's capacity to sustain and improve the health status and functioning of its Medicare population. The *Cohort IV Baseline* results are intended to assist M+COs and QIOs in identifying areas requiring potential improvement. The overall goals of HOS are to help beneficiaries make informed health care choices and to promote quality improvement based on competition.

In 2001, CMS required all M+COs and continuing cost contractors, Program of All-Inclusive Care for the Elderly (PACE) plans, Social HMOs, Medicare Choices, and DOD Subvention Demonstration Plans with contracts in place on or before January 1, 2000 to participate in the *Cohort IV Baseline* survey.

¹ HEDIS[®] is a registered trademark of the National Committee for Quality Assurance.

² National Committee for Quality Assurance. *HEDIS[®] 3.0/1998, Volume 6: Health of Seniors Survey Manual*. Washington DC: NCQA Publication, 1998.

For further background on the Medicare HOS, please refer to the *Cohort IV Baseline Report*. Additionally, the Medicare HOS Information and Technical Support Telephone Line (1-888-880-0077), as well as the HOS e-mail address (azpro.hos@sdps.org) are available to provide technical assistance.

MEDICARE HEALTH OUTCOMES SURVEY INSTRUMENT

The HOS instrument consists of the SF-36[®] Health Survey^{3,4} and additional questions, including those used for case mix and risk adjustment purposes.

SF-36[®] Health Survey

The SF-36[®] is a multi-purpose, short-form health survey with only 36 questions, which yields physical and mental health summary measures. It is a generic measure, as opposed to one that targets a specific age, disease, or treatment group. As documented in more than 2,500 publications, the SF-36[®] has proven useful in both general and specific populations, comparing the relative burden of diseases, differentiating the health benefits produced by a wide range of different treatments, and screening individual patients. The most complete information about the history and development of the SF-36[®], its psychometric evaluation, studies of reliability and validity, and normative data is available in two user's manuals.^{4,5}

Figure 1 on page 4 illustrates the taxonomy of items and concepts underlying the construction of the SF-36[®] physical and mental health summary measures. The SF-36[®] is scored from 0 to 100 points, with higher scores indicating better functioning on both the individual scales and the summary measures: Physical Component Summary Score (PCS) and Mental Component Summary Score (MCS). The HOS individual scale scores, as well as the PCS and MCS scores, have been normed to the values for the 1998 general US population, so that a score of fifty represents the national average for a given scale or summary score. In addition, the norm based score for the 1998 general US population has a standard deviation (SD) of ten points. It is important to note however, that the 1998 general population **elderly** norms used in the *Cohort IV Baseline Report* reflect a PCS mean score of 42.6 and an MCS mean score of 52.0.

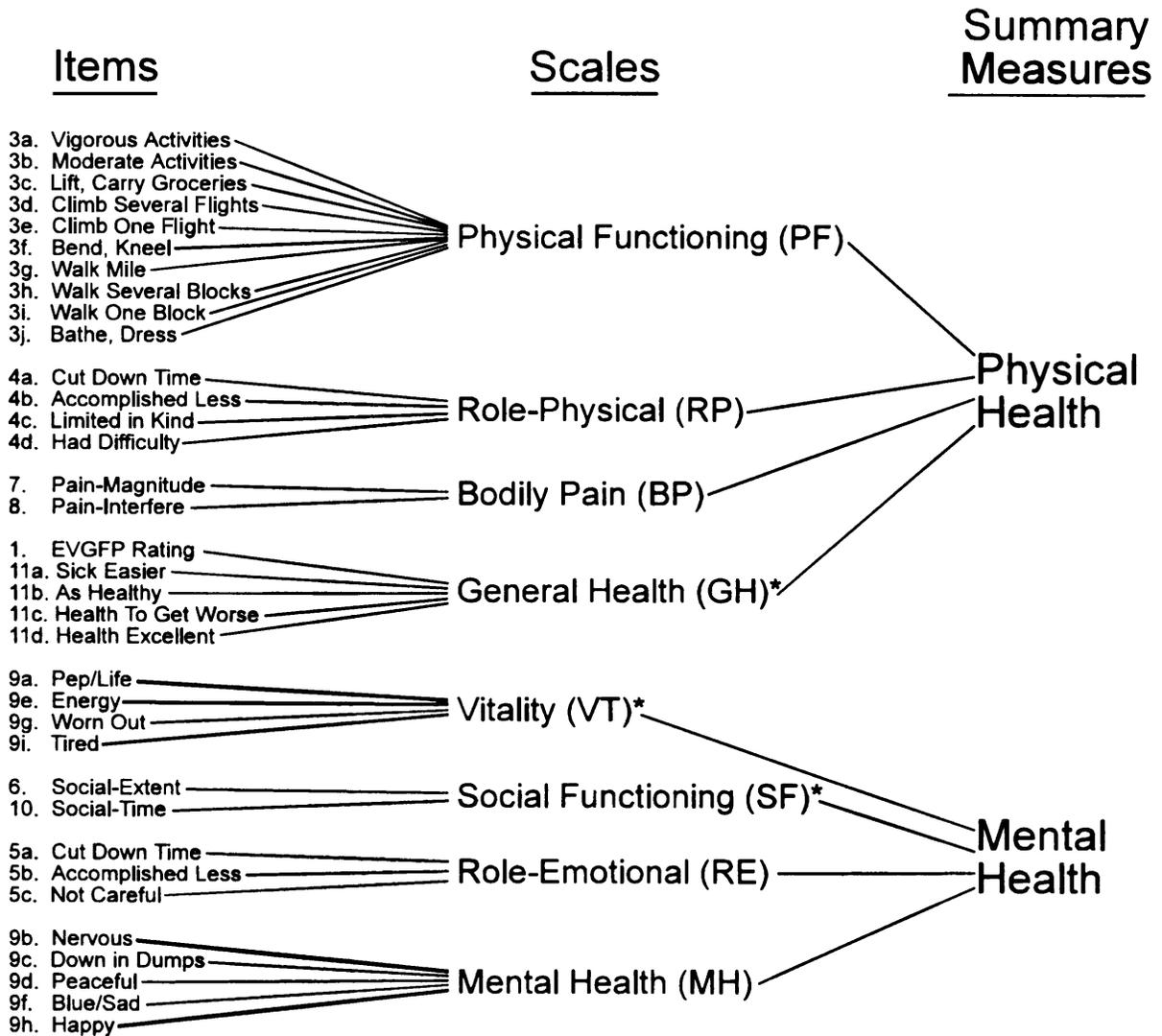
Please note that for prior HOS baseline data files, the means and standard deviations used in scoring PCS and MCS came from the 1990 National Survey of Functional Health Status. **Caution should be exercised when comparing PCS and MCS scores from the *Cohort IV Baseline* data file to all prior HOS baseline cohorts.**

³ SF-36[®] is a registered trademark of the Medical Outcomes Trust.

⁴ Ware JE, Snow KK, Kosinski M, Gandek B. *SF-36[®] Health Status Survey Manual and Interpretation Guide*. Boston, MA: The Health Institute, 1993.

⁵ Ware JE, Kosinski M. *SF-36[®] Physical and Mental Health Summary Scales: A Manual for Users of Version 1, Second Edition*. Lincoln, RI: QualityMetric, Incorporated, 2001.

FIGURE 1: SF-36[®] MEASUREMENT MODEL



* Significant correlation with other summary measure.

Source: Ware JE, Kosinski M, Keller SD. SF-36[®] Physical and Mental Health Summary Scales: A User's Manual. Boston, MA: The Health Institute, 1994.

Case Mix Adjustment Questions

The HOS instrument also includes questions on demographics, chronic medical conditions, and activities of daily living (ADLs). Demographics include questions on beneficiary age, gender, race, education, marital status, and income.

Thirteen chronic medical conditions are included in the questionnaire. These conditions are: hypertension; angina pectoris or coronary artery disease; congestive heart failure; myocardial infarction or heart attack; other heart conditions, such as heart valve defects or arrhythmias; stroke; emphysema, asthma, or chronic obstructive pulmonary disease; inflammatory bowel disease, including Crohn's disease and ulcerative colitis; arthritis of the hip or knee; arthritis of the hand or wrist; sciatica; diabetes, hyperglycemia, or glycosuria; and any cancer (other than skin cancer).

Six ADLs were included in the HOS survey to determine self reported difficulty with performance of daily tasks. Activities included bathing, dressing, eating, getting in or out of chairs, walking, and using the toilet.

For further information on the Medicare HOS instrument, please refer to the *Cohort IV Baseline Report* and Volume 6 of the 2001 HEDIS[®] Manual.⁶

⁶National Committee for Quality Assurance. *HEDIS[®] 2001, Volume 6: Specifications for the Medicare Health Outcomes Survey Manual*. Washington DC: NCQA Publication, 2001.

Methodology

SAMPLING METHODOLOGY

The HOS measure was administered to a randomly selected sample of individuals at baseline from each M+CO. The sampling methodology is dependent upon the plan's population. For M+COs with Medicare populations of more than 1,000 members, a simple random sample of 1,000 members was selected for the baseline survey. In those M+COs with 2,000 or more members, members who responded to the *Cohort III Baseline* survey were excluded from the *Cohort IV Baseline* sample. For M+COs with populations of 1,000 members or less, all eligible members were included in the sample for the baseline survey. Members were defined as eligible if they were continuously enrolled for at least 6 months and did not have End Stage Renal Disease (ESRD).

DISTRIBUTION OF THE SAMPLE

The 2001 *Cohort IV Baseline* Medicare HOS included a random sample of 190,523 beneficiaries, including both the aged and disabled, from 197 managed care plans. Of the 190,523 individuals sampled, 6,041 were determined to be invalid members during the survey administration. Invalid members of the sample meet one of the following criteria: deceased; not enrolled in the M+CO; have an incorrect address and phone number; or have a language barrier. The removal of the invalid members from the total sample yields an eligible sample of 184,482. This sample is referred to as the *Cohort IV Baseline eligible sample*. Of the 184,482 beneficiaries in the eligible sample, 68.4% (126,255) returned a completed baseline survey. For the purposes of the *Cohort IV Baseline* Report, a completed survey is defined as one that could be used to calculate a PCS and/or MCS score.

The 184,482 members of the *Cohort IV Baseline eligible sample* included 171,870 seniors (age 65 or older). Of the 171,870 eligible seniors sampled, 118,276 completed the baseline survey. This group of seniors comprises the *Cohort IV Baseline analytic sample*. The analytic sample is the focus of all analyses within the *Cohort IV Baseline* Report.

SCORING SF-36[®] PHYSICAL AND MENTAL HEALTH SUMMARY MEASURES

Physical and mental health status are estimated, respectively, using the PCS and MCS scoring algorithms recommended by the developers of the SF-36[®] Health Survey, as documented in detail elsewhere.⁷ Briefly, these norm-based algorithms yield favorably scored (i.e., higher is better) scales that have a mean of 50 and a standard deviation of 10 in the general US population. For PCS, very high scores indicate no physical limitations; disabilities or decline in well being; high energy level; and a rating of health as “excellent.” For MCS, very high scores indicate frequent positive affect, absence of psychological distress, and no limitations in usual social and role activities due to emotional problems.

So that population norms would be current, in relation to the timing of the first HOS cohort survey, the means and standard deviations used in scoring PCS and MCS for the *Cohort IV Baseline* data file came from the 1998 National Survey of Functional Health Status. So that PCS and MCS scores would have the same interpretation in the HOS as in previous studies, the weights (i.e., component scoring coefficients) used in aggregating the eight scales to score each of those summaries are the original standardized weights recommended by the developers.⁸ These weights, which have been used in more than 100 published studies reporting results for the PCS and MCS summary measures, have consistently yielded reliable and valid scores in both general and elderly populations. Given this consistency and reliability, the published interpretation guidelines are applicable to the HOS.

Please note that for prior HOS baseline data files, the means and standard deviations used in scoring PCS and MCS came from the 1990 National Survey of Functional Health Status. **Caution should be exercised when comparing PCS and MCS scores from the *Cohort IV Baseline* data file to all prior HOS baseline cohorts.**

The HOS is among the first large scale surveys to take advantage of improved algorithms for scoring the PCS and MCS summary measures for respondents with missing data. Most previous studies have used the “half scale” rule for imputing scale scores for those with missing data. This solution, which was developed during the Health Insurance Experiment more than 20 years ago, is widely used in health status research.⁹ However, the “half scale” approach has several disadvantages, including: being applicable only to those with at least half of the items answered for each of the eight scales; introducing a bias in score estimates because answered items are simply averaged in estimating missing items; and failing to provide an estimation strategy for PCS and MCS for those with a missing scale score.

The improved scoring algorithms use the missing data estimation (MDE) utility. The MDE

⁷ Ware JE, Kosinski M. *SF-36[®] Physical and Mental Health Summary Scales: A Manual for Users of Version 1, Second Edition*. Lincoln, RI: QualityMetric, 2001.

⁸ Ware JE, Kosinski M. *SF-36[®] Physical and Mental Health Summary Scales: A User's Manual*. Lincoln, RI: QualityMetric, 2001.

⁹ Ware JE, Brook RH, Davies-Avery A, Williams K, Stewart AL, Rogers WH, et al. *Model of Health and Methodology*. Santa Monica, CA: RAND Corporation, 1980; R-1987/1-HEW. (*Conceptualization and Measurement of Health for Adults in the Health Insurance Study; vol. 1*).

scoring utility, which was validated using item response theory, calculates an unbiased score as long as at least one item is answered within each scale. Further, the MDE software uses regression methods to score PCS and MCS for those with one scale missing. As documented elsewhere, the MDE scoring algorithms have been evaluated in the 1998 general US population and in the HOS.¹⁰ In the HOS *Cohort IV Baseline* sample, the MDE software calculated summary scores for an additional 5,047 (2.6%) study participants. These scores would have previously been lost at baseline due to missing data. *Please note, the MDE scoring utility does not output scale level results; therefore, scale scores were not included in this data file.*

CASE MIX ADJUSTMENT

Of the 171,870 eligible seniors sampled, 118,276 had a calculatable PCS and/or MCS score. Linear regression techniques were used to case mix adjust these scores for each beneficiary. In brief, models used to adjust PCS and MCS scores included variables to control for differences in demographic and socioeconomic characteristics, chronic medical conditions, and HOS study design variables. Demographic and socioeconomic variables included age, gender, race, education, marital status, and income. Chronic medical conditions were measured with a checklist of 13 medical conditions. HOS study design variables included who completed the survey, the mode of survey administration, CMS region, and the survey vendor. The case mix adjustment of PCS and MCS scores was limited to those beneficiaries with complete data for all covariates included in the model. The following table, Table 1 on page 9, describes the covariates used in the case mix adjustment of the SF-36[®] measures.

For further information on the Medicare HOS methodology, please refer to the *Cohort IV Baseline Report*.

¹⁰ Kosinski MK, Bayliss M, Bjorner JB, Ware JE. *Improving Estimates of SF-36[®] Health Survey Scores for Respondents in Missing Data*. Medical Outcomes Trust Monitor, Fall 2000; 5 (1): 8-10.

TABLE 1
COVARIATES USED IN THE CASE MIX ADJUSTMENT OF
SF-36® SUMMARY MEASURES

DEMOGRAPHICS	<ul style="list-style-type: none"> ➤ Age (Continuous) ➤ Gender (Male or Female) ➤ Race (White, Black, Other Minority) ➤ Education ➤ Marital Status ➤ Income
CHRONIC MEDICAL CONDITIONS	<ul style="list-style-type: none"> ➤ Hypertension or high blood pressure ➤ Angina pectoris or coronary artery disease ➤ Congestive heart failure ➤ Myocardial infarction or heart attack ➤ Other heart conditions, such as problems with heart valves or arrhythmias ➤ Stroke ➤ Emphysema, or asthma, or COPD (Chronic Obstructive Pulmonary Disease) ➤ Crohn's disease, ulcerative colitis, or inflammatory bowel disease ➤ Arthritis of the hip or knee ➤ Arthritis of the hand or wrist ➤ Sciatica ➤ Diabetes, high blood sugar, or sugar in the urine ➤ Any cancer (other than skin cancer)
HOS STUDY DESIGN VARIABLES	<ul style="list-style-type: none"> ➤ Who Completed Survey (Self or Other) ➤ Mode of Survey Administration (Mail or Telephone) ➤ CMS Plan Region ➤ Survey Vendor

Baseline Electronic Data File Characteristics

Each QIO level *Cohort IV Baseline* Electronic Data File was derived from a sample of 190,523 Medicare beneficiaries. There are a total of 241 variables in the *Cohort IV Baseline* Electronic Data File. A detailed summary of data cleaning, editing and scoring processes are included in Appendix A. A detailed list of these variables is included in Appendix B. The following is an overview of the variables included in the data file.

To facilitate usage of the HOS data file, the following variables have been added to the data file in 2001.

In place of one address field for the beneficiary address, there are six address fields (*C4ADDRS1* – *C4ADDRS6*). The beneficiary's Social Security Number (*C4SSN*) has been added to the file. A protocol identifier flag (*C4PROTIF*) indicates the beneficiary's protocol status.

In place of one name field for the name of person completing the survey, there are three name fields, (*C4FNMCMP*, *C4MNMCMP*, *C4LNMCMP*) to accommodate a first, middle and last name. A variable (*C4CMPFLG*) was created to indicate if any of these three fields has data. For beneficiaries who are completing a *Cohort II Follow Up* survey and had a proxy at baseline, the member's response to Question 56 of the *Cohort II Baseline* survey (*C4RESPC2*) is included. The variable *C4PROXST* has been added to the file with an indicator of "9" for all records of the *Cohort IV Baseline* file.

For beneficiaries sampled for both the *Cohort IV Baseline* and *Cohort II Follow Up* surveys, the variable *C4DBLDTY* indicates the member's "double-duty" status. The wave (one or two) in which the data were submitted to NCQA by the vendors is indicated by the variable *C4ROUND*. The variable *C4VUCATI* indicates the vendor's unique Computer Assisted Telephone Interviewer (CATI) ID for surveys with a telephone (T) final disposition code other than T24 or T36.

Two variables were created for the data file that give categories for the race group (*C4RACEGP*) and age group (*C4AGEGP2*) of beneficiaries. In addition, a new market area variable (*C4RPT_MA*) contains information about the market area designation that corresponds to the market area in the reporting unit variable (*C4RPTUNT*).

Below please find a general description of variables included in the *Cohort IV Baseline* data file.

PLAN LEVEL VARIABLES (VARIABLES 1-11)

Data from this section are taken from the header of the *Cohort IV Baseline* M+CO data file. This section includes information about the M+CO contract number, plan name, market area, and vendor.

BENEFICIARY LEVEL VARIABLES (VARIABLES 12-39)

This section contains confidential beneficiary level data. Data from this section are taken from the member level record of the baseline data file abstracted from the CMS Medicare Enrollment Database. The Health Insurance Claim (HIC) Number (*HICNUM*) is a unique identifier used to identify each beneficiary in the data file. This variable is the beneficiary level unit of analysis for the *Cohort IV Baseline* Report. Beneficiary addresses (mailing address, county, state, and zip code), race, gender, and reason for entitlement are included in this section.

SURVEY LEVEL VARIABLES (VARIABLES 40-140)

This section contains survey information from the 57 questions comprising the HOS instrument. The information presented in this section represents each beneficiary's actual answers to the specific questions. These data include beneficiary responses to questions pertaining to the SF-36[®], health status indicators, chronic medical conditions, depression, ADLs, proxy status, and demographics. The SF-36[®] (questions one and three through eleven in the HOS instrument) is used to obtain physical and mental health summary measures. The beneficiary level data can potentially be used by individuals to perform quality improvement initiatives.

PLAN, VENDOR, AND SURVEY STATISTICS (141-176)

The variables in this section include vendor derived information about the beneficiary status such as survey language, survey date, and the round in which the survey was obtained. Invalid survey indicator, and type of survey completed, such as mail or telephone, are included. In addition, there is plan information obtained from the CMS **June 2001** Monthly Report of Medicare Coordinated Care Health Plans (<http://www.cms.hhs.gov/statistics/monthly.asp>). An important variable in this group is *C4PLANST*, which is the two letter state abbreviation as listed in the CMS June 2001 Monthly Report. This was the state level unit of analysis for the *Cohort IV Baseline* Report.

ANALYTIC VARIABLES (177-205)

This section includes variables created for analytic purposes. The *Cohort IV Baseline analytic sample* in the M+CO data file is indicated by the variable *C4ANALYT* and is limited to beneficiaries who are 65 or older and who had a calculatable PCS and/or MCS score. The variable *C4AGE* has been calculated as an exact age by the following formula in SAS code where *C4EDOB* was the elapsed date of birth and *C4ESVDAT* was the elapsed date of the survey:

$$C4AGE = \text{yrdif}(C4EDOB, C4ESVDAT, 'act/act')$$

When the survey date was not available in the data file an intermediate survey date of July 1, 2001 was used for the elapsed survey date. The reporting unit variable (*C4RPTUNT*) was the plan unit of analysis for the *Cohort IV Baseline Report*. A unique beneficiary ID (*C4PATID*) is also included. This variable is a unique number assigned to each beneficiary in the *Cohort IV Baseline* sample (1 to 190,523).

In the *Cohort I, II, and III Baseline* data files previously distributed, selected variables were only calculated for the analytic population. However, to facilitate analysis of populations not included in the *Cohort IV Baseline analytic sample*, these variables were calculated for all members of the sample with available data. These variables include: the number of chronic medical conditions (*C4COMO*), the categories for number of chronic medical conditions (*C4COMOCT*), an indicator for a positive depression screen (*C4DEPRSS*), and race group categories (*C4RACEGP*). The age group categories for the variable *C4AGEGP2* were limited to those 65 and older.

SF-36[®] SUMMARY MEASURES (VARIABLES 206-241)

The variables that are presented in this section include the unadjusted PCS (*C4BASPCS*) and MCS (*C4BASMCS*) scores. These scores were generated by the MDE software. The HOS individual scale scores, as well as the PCS and MCS scores, have been normed to the values for the 1998 general US population, so that a score of fifty represents the national average for a given scale or summary score. *Please note, the MDE scoring utility does not output scale level results; therefore, scale scores were not included in this data file.* In addition, the case mix adjusted PCS scores (*C4ADJPCS*) and MCS scores (*C4ADJMCS*) are included. The unadjusted and adjusted plan, state, and national means and standard deviations for PCS and MCS are included at the beneficiary level (*C4PCMNTU – C4MCSDSA*).

For further details on the case mix adjustment variables and the data analysis, please refer to Appendix B and the *Cohort IV Baseline Report*. Additionally, the Medicare HOS Information and Technical Support Telephone Line (1-888-880-0077), as well as the HOS e-mail address (azpro.hos@sdps.org) are available to provide technical assistance.

PACE Discussion

The HOS *Cohort IV Baseline* sample included beneficiaries enrolled PACE plans. There were 3,943 members sampled from 17 PACE plans. There are 242 variables in the baseline PACE data file. Since the PACE plans differ significantly from the M+CO plans, their data were analyzed separately and have been forwarded in a separate data file. **Even though the structures of the M+CO and PACE data files are similar (see Appendix B), it is not recommended that the data files be combined.**

The PACE data file was prepared in a similar manner to the M+CO data file, however some special exceptions were allowed. The *Cohort IV Baseline analytic sample* in the M+CO data file is indicated by the variable *C4ANALYT* and is limited to beneficiaries who are 65 or older and who had a calculatable PCS and/or MCS score. For *Cohorts II*, and *III Baseline* a unique analytic variable was created for the PACE beneficiaries due to missing information, (i.e., age or reason for entitlement). As a result of this missing information, the criteria were waived for PACE beneficiaries. In order to be consistent with the previous cohort data files, a special PACE specific analytic indicator *C4ANAL_P* was created which indicates those PACE beneficiaries who are 65 or older and had a calculatable PCS and/or MCS score.

For PACE beneficiaries, the unadjusted and adjusted plan, PACE state total, and overall PACE total means and standard deviations for PCS and MCS are included (*C4PCMNTU – C4MCSDSA*).

When the general plan characteristics information was downloaded from the CMS **June 2001** monthly report of Medicare Coordinated Care Health Plans from the following website: (<http://www.cms.hhs.gov/statistics/monthly.asp>) PACE plans were not included in the monthly report. Therefore, the variables *CMONRPT – C4PLNDCT* are missing for all PACE plans. The variable, which is used to indicate plan state (*C4PLANST*), is obtained from this CMS monthly report. Because this variable was missing for all PACE plans, beneficiary address information was used to ascertain plan to state relationship, and a new variable (*C4PLST_P*) was created only for the PACE plans. This new state variable was also used to assign CMS region status (*C4PLRG_P*).

The PACE data file does not contain the variable *C4DBLDTY* since the PACE beneficiaries are sampled only for the baseline survey.

For further details on the PACE data file, please refer to Appendix B and the *Cohort IV Baseline Report*. Additionally, the Medicare HOS Information and Technical Support Telephone Line (1-888-880-0077), as well as the HOS e-mail address (azpro.hos@sdps.org) are available to provide technical assistance.

Appendix A

Quality Assurance of the Data

DATA RECEIPT

HOS *Cohort IV Baseline* data were transmitted to Health Services Advisory Group (HSAG) from the National Committee for Quality Assurance (NCQA). The data were transmitted on CD-ROM containing individual ACSII flat files for each participating plan. These plans included all Medicare + Choice Organizations (M+COs) and continuing cost contractors, Program of All-Inclusive Care for the Elderly (PACE) plans, Social HMOs, Medicare Choices, and DOD Subvention Demonstration Plans with contracts in place on or before January 1, 2000. In total, 386 individual files were submitted containing 190,523 baseline records and 3,943 PACE records. The ASCII flat files contained plan, beneficiary, and survey information as specified in Volume 6 of the 2001 HEDIS[®] Manual (also referred to as the 2001 HOS Manual) and the 2001 NCQA Quality Assurance Plan (also referred to as the 2001 QAP).^{1,2}

DATA CLEANING AND EDITING

Once all of the individual files were successfully imported using SAS software and saved as unique SAS data sets, they were appended to form a single HOS data file. To verify the presence of unique beneficiaries in the HOS data file, the file was examined for duplicate Health Insurance Claim (HIC) Numbers. Data consistency checks were performed by reviewing the entire HOS data set for out of range values. All dates contained within the data file were verified to correspond to the appropriate range. Frequency distributions of all categorical variables as well as cross tabulations by vendor were performed to identify both out of range values and data shifts in value assignment. The cross tabulations were performed using the entire HOS data file and also specified subsets of the data file.

An additional consistency check was performed which examined skip pattern violations. In many records, beneficiaries failed to correctly follow the skip patterns contained within the survey; however, no changes were made to any of the responses. **Caution should be exercised when examining data which utilizes a skip pattern.**

Several inconsistencies pertaining to survey date, survey language, and Spanish protocol identifier flag were identified. An example of this type of inconsistency is a record with a survey disposition code which indicates that the survey was a nonresponse/refusal (“T32”), but the survey date is not equal to “99999999”. As identified, the errors were forwarded to NCQA, and

¹ National Committee for Quality Assurance. *HEDIS[®] 2001, Volume 6: Specifications for the Medicare Health Outcomes Survey*. Washington DC: NCQA Publication, 2001.

² National Committee for Quality Assurance. *NCQA Quality Assurance Plan (QAP): 2001 Medicare Health Outcomes Survey Cohort 4 Baseline Administration and Cohort 2 Follow Up Administration*. Washington DC: NCQA Publication, revised June 8, 2001.

when appropriate corrected information was incorporated into the data file. **Caution should be exercised when examining survey disposition, round number, survey date, survey language, and Spanish protocol flag variables during analysis.**

After the HOS data file was cleaned and edited, additional variables were added to the file. An indicator for the beneficiaries of the PACE plans (*C4PACE*) was added. To be consistent with the variables included in the *Cohort II* and *III Baseline* data files, an EverCare indicator (*C4EVER*) is created; however no EverCare plans were included in Round 4 of HOS. In addition, state names (*C4STNAME*) and abbreviations (*C4STABV*) were incorporated which decode the Social Security Administration's (SSA) numeric state values from the member level record information.

Some demographic variables were contained in both the member level data provided by CMS and the respondent provided survey data. When inconsistencies were found between these two data sources indicator variables were created. Indicators were created for the following: mismatched year of birth (*C4BDBRTH*), mismatched race (*C4BDRACE*), mismatched gender (*C4BDGNDR*), and female beneficiaries reporting current treatment for prostate cancer (*C4BDPRST*).

Plan specific variables which were added to the file included number of ineligible beneficiaries, sample size, total number of completed surveys, number completed by mail, number completed by telephone, overall response rate, mail response rate, and telephone response rate. All date variables contained in the data file were converted to SAS date format (elapsed date variables) to facilitate the calculation of duration of enrollment (*C4ENRDUR*) and age (*C4AGE*), which were then incorporated into the data file.

In addition to those variables listed above, general plan characteristics information were downloaded from the CMS **June 2001** Monthly Report of Medicare Coordinated Care Health Plans (<http://www.cms.hhs.gov/statistics/monthly.asp>). The statistics from the monthly report were incorporated into the *Cohort IV Baseline* data file and included the following plan specific variables: type; model; population; description; CMS region; state; tax status; and contract start date. Duration of plan contract in years (*C4PLDUR*) was calculated and incorporated into the data file.

Upon completion of the HOS data editing and cleaning process, the final data set was produced. This final data set served as the source data set for the baseline analysis.

For further information on the quality assurance of the data, please refer to the *Cohort IV Baseline* Report. Additionally, the Medicare HOS Information and Technical Support Telephone Line (1-888-880-0077), as well as the HOS e-mail address (azpro.hos@sdps.org) are available to provide technical assistance.

Appendix B

Baseline Electronic Data File Specifications

The following table describes the file layout by position for the *Cohort IV Baseline* Electronic Data File. There are a total of 241 variables in the data file. The file is a SAS data file and was generated using the SAS version 6 engine.

VAR #	FIELD NAME	VAR TYPE	LENGTH/FORMAT	FIELD POSITION	LABEL DESCRIPTION
1	C4RECID	Char	1	0	C4 Record ID
2	C4RPTYR	Num	8	1	C4 Report Year
3	C4CNTRNM	Char	5	9	C4 Contract Number
4	C4PLANNM	Char	50	14	C4 Plan Name
5	C4LNBUSS	Char	1	64	C4 Line of Business
6	C4MODEL	Char	1	65	C4 Model Type
7	C4MARKET	Char	1	66	C4 Market Area
8	C4VENDOR	Num	4	67	C4 Vendor 1 = DSS 2 = GHS 3 = Solucient 4 = Market Facts
9	C4HTHID	Char	7	71	C4 NCQA Healthcare Organization ID
10	C4SUBID	Char	6	78	C4 NCQA Submission ID
11	C4SPECID	Char	10	84	C4 NCQA Special Area ID
12	HICNUM	Char	12	94	HIC Number
13	C4PLANID	Char	6	106	C4 Plan ID
14	C4FNAME	Char	15	112	C4 First Name
15	C4MI	Char	1	127	C4 Middle Initial
16	C4LNAME	Char	12	128	C4 Last Name
17	C4ADDRS1	Char	22	140	C4 Address Line 1
18	C4ADDRS2	Char	22	162	C4 Address Line 2
19	C4ADDRS3	Char	22	184	C4 Address Line 3
20	C4ADDRS4	Char	22	206	C4 Address Line 4
21	C4ADDRS5	Char	22	228	C4 Address Line 5
22	C4ADDRS6	Char	22	250	C4 Address Line 6
23	C4STATE	Char	2	272	C4 State SSA Code
24	C4COUNTY	Char	3	274	C4 County SSA Code
25	C4ZIPCOD	Char	5	277	C4 Zip Code
26	C4SSN	Char	9	282	C4 Social Security Number
27	C4RACE	Num	4	291	C4 Race (CMS)
28	C4GENDER	Num	4	295	C4 Gender (CMS)
29	C4DOB	Char	8	299	C4 Date of Birth
30	C4DOD	Char	8	307	C4 Date of Death

VAR #	FIELD NAME	VAR TYPE	LENGTH/FORMAT	FIELD POSITION	LABEL DESCRIPTION
31	C4ACCRDT	Char	8	315	C4 Accretion Date
32	C4TERMDT	Char	8	323	C4 Termination Date
33	C4ESRDID	Num	4	331	C4 ESRD Indicator
34	C4INSTUT	Num	4	335	C4 Institutional Status
35	C4HOSPIC	Num	4	339	C4 Hospice Status
36	C4MEDICD	Num	4	343	C4 Medicaid Status
37	C4RSENT	Num	4	347	C4 Reason for Entitlement
38	C4PROTIF	Num	8	351	C4 Protocol Identifier Flag
39	C4RESPC2	Char	50	359	C4 Respondent to Cohort 2 Survey
40	C4GENHHTH	Num	4	409	C4 Q1 General Health Question
41	C4HTHTRN	Num	4	413	C4 Q2 Health Transition Question
42	C4VIGACT	Num	4	417	C4 Q3a Vigorous Activities
43	C4MODACT	Num	4	421	C4 Q3b Moderate Activities
44	C4LIFT	Num	4	425	C4 Q3c Lift carry groceries
45	C4CLMBSV	Num	4	429	C4 Q3d Climb several flights
46	C4CLMBON	Num	4	433	C4 Q3e Climb one flight
47	C4BEND	Num	4	437	C4 Q3f Bending
48	C4WLKMI	Num	4	441	C4 Q3g Walk > 1 mile
49	C4WLKBKS	Num	4	445	C4 Q3h Walk several blocks
50	C4WLK1BK	Num	4	449	C4 Q3i Walk 1 block
51	C4BATHDR	Num	4	453	C4 Q3j Bathing or dressing
52	C4PCUTTM	Num	4	457	C4 Q4a Phys-Cut down time
53	C4PACMPL	Num	4	461	C4 Q4b Phys-Accomplish less
54	C4PLMTKW	Num	4	465	C4 Q4c Lmtd in Kind of work
55	C4PDIFWK	Num	4	469	C4 Q4d Difficulty w/work
56	C4ECUTTM	Num	4	473	C4 Q5a Emotnl-cut down time
57	C4EACMPL	Num	4	477	C4 Q5b Emotnl-Accomplished less
58	C4ENTCRF	Num	4	481	C4 Q5c Emotnl-Not careful
59	C4SOCLMT	Num	4	485	C4 Q6 Social-extent
60	C4PNMAGT	Num	4	489	C4 Q7 Pain-Magnitude
61	C4PNINTF	Num	4	493	C4 Q8 Pain-Interfere Work
62	C4FULPEP	Num	4	497	C4 Q9a Pep/Life
63	C4NERVS	Num	4	501	C4 Q9b Nervous
64	C4DNDMPS	Num	4	505	C4 Q9c Down in Dumps
65	C4PCEFUL	Num	4	509	C4 Q9d Peaceful
66	C4ENERGY	Num	4	513	C4 Q9e Energy
67	C4BLSAD	Num	4	517	C4 Q9f Blue/Sad
68	C4WRNOUT	Num	4	521	C4 Q9g Worn Out
69	C4HAPPY	Num	4	525	C4 Q9h Happy
70	C4TIRED	Num	4	529	C4 Q9i Tired

VAR #	FIELD NAME	VAR TYPE	LENGTH/ FORMAT	FIELD POSITION	LABEL DESCRIPTION
71	C4SCLACT	Num	4	533	C4 Q10 Social-time
72	C4SCKESY	Num	4	537	C4 Q11a Sick Easier
73	C4ASHLTH	Num	4	541	C4 Q11b As Healthy
74	C4HTHWSE	Num	4	545	C4 Q11c Expct Hth to get worse
75	C4HTHEXT	Num	4	549	C4 Q11d Health Excellent
76	C4DIFBTH	Num	4	553	C4 Q12a Difclty Bathing
77	C4DIFDRS	Num	4	557	C4 Q12b Difclty Dressing
78	C4DIFEAT	Num	4	561	C4 Q12c Difclty Eating
79	C4DIFCHR	Num	4	565	C4 Q12d Difclty Getng out chair
80	C4DIFWLK	Num	4	569	C4 Q12e Difclty Walking
81	C4DIFTOL	Num	4	573	C4 Q12f Difclty Using Toilet
82	C4CPNEXR	Num	4	577	C4 Q13a Chest Pain-Exercise
83	C4CPNRST	Num	4	581	C4 Q13b Chest Pain-Resting
84	C4SOBFLT	Num	4	585	C4 Q14a Short Brth lying flat
85	C4SOBSIT	Num	4	589	C4 Q14b Short Brth sittng/restng
86	C4SOBWLK	Num	4	593	C4 Q14c Short Brth Walk<1 block
87	C4SOBSTR	Num	4	597	C4 Q14d Short Brth Climb 1 flight stairs
88	C4NMBFET	Num	4	601	C4 Q15a Numbness Feet
89	C4ANKSWL	Num	4	605	C4 Q15b Ankles Swell
90	C4TINGFT	Num	4	609	C4 Q15c Tingling Burning feet
91	C4DECSNS	Num	4	613	C4 Q15d Decreased feel hot/cold
92	C4DECHEL	Num	4	617	C4 Q15e Sores that dont heal
93	C4PARLYS	Num	4	621	C4 Q16a Ever had paralysis
94	C4LSTTLK	Num	4	625	C4 Q16b Ever lost ability to talk
95	C4RDNEWP	Num	4	629	C4 Q17 See to read newspaper
96	C4HRMOST	Num	4	633	C4 Q18 Hear most things
97	C4ACDING	Num	4	637	C4 Q19 Acid Indigestion
98	C4CTRURN	Num	4	641	C4 Q20 Difclty cntrl urination
99	C4HIGHBP	Num	4	645	C4 Q21 HTN or High BP
100	C4ANGCAD	Num	4	649	C4 Q22 Angina or CAD
101	C4CHF	Num	4	653	C4 Q23 CHF
102	C4AMI	Num	4	657	C4 Q24 AMI
103	C4OTHHRT	Num	4	661	C4 Q25 Other Heart cond.
104	C4STROKE	Num	4	665	C4 Q26 Stroke
105	C4COPD E	Num	4	669	C4 Q27 Emphy. Asthma COPD
106	C4GI ETC	Num	4	673	C4 Q28 GI Problems
107	C4ATHHIP	Num	4	677	C4 Q29 Arthritis hip or knee
108	C4ATHHAN	Num	4	681	C4 Q30 Arthritis hand or wrist
109	C4SCIATC	Num	4	685	C4 Q31 Sciatica
110	C4DIABET	Num	4	689	C4 Q32 Diabetes
111	C4ANYCAN	Num	4	693	C4 Q33 Any Cancer
112	C4ARTHPN	Num	4	697	C4 Q34 Level arth pain
113	C4COLNCA	Num	4	701	C4 Q35a Colon Cancer
114	C4LUNGCA	Num	4	705	C4 Q35b Lung Cancer

VAR #	FIELD NAME	VAR TYPE	LENGTH/FORMAT	FIELD POSITION	LABEL DESCRIPTION
115	C4BRSTCA	Num	4	709	C4 Q35c Breast Cancer
116	C4PROSCA	Num	4	713	C4 Q35d Prostate Cancer
117	C4BACKPN	Num	4	717	C4 Q36 Back Pain Interferes
118	C4NUMBLG	Num	4	721	C4 Q37 Numbness in leg
119	C4FELTSD	Num	4	725	C4 Q38 Sad/Blue in past 2 wks
120	C4DEPMCH	Num	4	729	C4 Q39 Depressed in past yr
121	C4DEP2YR	Num	4	733	C4 Q40 Two + years depressed
122	C4CMPHTH	Num	4	737	C4 Q41 Hth compared to peers
123	C4SMK100	Num	4	741	C4 Q42 Ever smoke 100 cigs
124	C4SMKFRQ	Num	4	745	C4 Q43 Curnt smoke freq
125	C4DRSQT	Num	4	749	C4 Q44 Time since Quit smoke
126	C4QSMKAD	Num	4	753	C4 Q45 Freq antismoke advice
127	C4BRTHYR	Char	4	757	C4 Q46 Survey-Birth Yr
128	C4SV_GND	Num	4	761	C4 Q47 Survey-Gender
129	C4HISPAN	Num	4	765	C4 Q48 Hispanic Y/N
130	C4SV_RAC	Num	4	769	C4 Q49 Survey-Race
131	C4MARITL	Num	4	773	C4 Q50 Marital Status
132	C4EDUC	Num	4	777	C4 Q51 Education Level
133	C4HMOWN	Num	4	781	C4 Q52 Homeowner Status
134	C4RTRCOM	Num	4	785	C4 Q53 Retirement Community
135	C4MDSVPV	Num	4	789	C4 Q54 Medical Svcs Provided
136	C4WHOCMP	Num	4	793	C4 Q55 Who completed Survey
137	C4FNMCMC	Char	15	797	C4 Q56 First Name Person comp
138	C4MNMCMC	Char	15	812	C4 Q56 Middle Name Person comp
139	C4LNMCMC	Char	20	827	C4 Q56 Last Name Person comp
140	C4HHINC	Num	4	847	C4 Q57 Household Income
141	C4SRVDSP	Char	3	851	C4 Survey Disposition
142	C4RNDNUM	Char	2	854	C4 Round Survey Obtained
143	C4SVLANG	Num	4	856	C4 Survey Language 1 = English 2 = Spanish 3 = Not Applicable
144	C4SVDATE	Char	8	860	C4 Date Survey Completed
145	C4PROXST	Num	4	868	C4 Proxy Status
146	C4DBLDTY	Num	8	872	C4 Double Duty (in C4B and R2) 0 = No 1 = Yes
147	C4ROUND	Num	8	880	C4 Round Data Submitted (1 or 2)
148	C4CMPFLG	Num	8	888	C4 Name Provided (Q56) flag
149	C4VUCATI	Char	8	896	C4 Vendor Unique CATI ID
150	C4MCONUM	Num	4	904	C4 MCO Provide Phone Num
151	C4SPANFL	Num	4	908	C4 Spanish Protocol Flag
152	C4PACE	Num	8	912	C4 Indicator PACE Plan
153	C4EVER	Num	8	920	C4 Indicator EVERCARE Plan
154	C4INVSrv	Num	8	928	C4 Invalid Survey

VAR #	FIELD NAME	VAR TYPE	LENGTH/ FORMAT	FIELD POSITION	LABEL DESCRIPTION
155	C4PCTCMP	Num	8	936	C4 Percent of Survey Completed
156	C4CMPSRV	Num	8	944	C4 Completed Survey (80% Criteria) 0 = No 1 = Yes
157	C4INVALID	Num	8	952	C4 Ineligible by Plan
158	C4SAMPLE	Num	8	960	C4 Sample Size by Plan
159	C4RSPNUM	Num	8	968	C4 Completed Surveys by Plan
160	C4MALNUM	Num	8	976	C4 Mail Completed Surveys by Plan
161	C4TELNUM	Num	8	984	C4 Tele Completed Surveys by Plan
162	C4RSPDEN	Num	8	992	C4 Sample minus invalids by Plan
163	C4RSPON	Num	8	1000	C4 Response Rate by Plan
164	C4MALRR	Num	8	1008	C4 Mail Response Rate by Plan
165	C4TELRR	Num	8	1016	C4 Tele Response Rate by Plan
166	C4MONRPT	Char	8	1024	C4 Monthly Report - source CMS 06/01
167	C4PLTYPE	Char	3	1032	C4 Plan Type – source CMS 06/01
168	C4PLMODL	Char	5	1035	C4 Plan Model – source CMS 06/01
169	C4PLPOP	Char	8	1040	C4 Plan Population - source CMS 06/01
170	C4PLDESC	Char	5	1048	C4 Plan Description - source CMS 06/01
171	C4PLANST	Char	2	1053	C4 Plan State – source CMS 06/01
172	C4PLTXST	Char	4	1055	C4 Plan Tax Status - source CMS 06/01
173	C4PLREG	Num	8	1059	C4 Plan CMS Region - source CMS 06/01
174	C4PLSTDT	Num	8/ MMDDYY10	1067	C4 Contract Start Date -source CMS 06/01
175	C4PLDUR	Num	8	1075	C4 Duration of Plan Contract - Years
176	C4PLNDCT	Num	8	1083	C4 Duration of Plan Contract Categories 1 = Less than 1 Year 2 = 1 to 4.9 Years 3 = 5.0 to 9.9 Years 4 = 10 or more Years
177	C4PATID	Num	8	1091	C4 Anonymous Patient ID 1 to 190,523
178	C4PLAN	Num	8	1099	C4 Anonymous Plan ID 1 to 197
179	C4EDOB	Num	8/ MMDDYY10	1107	C4 Elapsed Date of Birth
180	C4AGE	Num	8	1115	C4 Age (Continuous)
181	C4AGEGRP	Num	8	1123	C4 Age Groups 1 = Less than 55 2 = 55 to 64 3 = 65 to 74 4 = 75 or older
182	C4AGEGP2	Num	8	1131	C4 Age Groups for Baseline Report 1 = C4AGE ≥ 65 and <70 2 = C4AGE ≥ 70 and <75 3 = C4AGE ≥ 75 and <80 4 = C4AGE ≥ 80 (where C4AGE ≥ 65)

VAR #	FIELD NAME	VAR TYPE	LENGTH/ FORMAT	FIELD POSITION	LABEL DESCRIPTION
183	C4RACEGP	Num	8	1139	C4 Race Groups for Baseline Report 1 = White 2 = Black 3 = Other (CMS data used – C4Race)
184	C4ANALYT	Num	8	1147	C4 Analytic Indicator 0 = No 1 = Yes (where C4AGE ≥ 65 and has PCS and/or MCS)
185	C4DEPRSS	Num	8	1155	C4 Positive Depression Screen (Yes to any Question 38, 39, or 40)
186	C4COMO	Num	8	1163	C4 Number of Comorbidis
187	C4COMOCT	Num	8	1171	C4 Comorbid Categories 1 = 0 or 1 Condition (s) 2 = 2 or 3 Conditions 3 = 4 or more Conditions
188	C4EDOD	Num	8/ MMDDYY10	1179	C4 Elapsed Date of Death
189	C4ETRMDT	Num	8/ MMDDYY10	1187	C4 Elapsed Date of Termination from Plan
190	C4EACRDT	Num	8/ MMDDYY10	1195	C4 Elapsed Date of Accretion into Plan
191	C4ESVDAT	Num	8/ MMDDYY10	1203	C4 Elapsed Date of Survey
192	C4EACLMT	Num	8/ MMDDYY10	1211	C4 Elapsed Date-Accretion Limit-3/01/01
193	C4ENRDUR	Num	8	1219	C4 Enrollment Duration (Months)
194	C4ENRCAT	Num	8	1227	C4 Enrollment Duration Categories 1 = Less than 6 Months 2 = 6 to 12 Months 3 = 13 to 36 Months 4 = 37 or more Months
195	C4BDCNUM	Num	8	1235	C4 Mismatched Contract Number 0 = No 1 = Yes (Plan ID versus Contract Number)
196	C4BDRACE	Num	8	1243	C4 Mismatched Race 0 = No 1 = Yes (CMS Data versus Survey Response)
197	C4BDBRTH	Num	8	1251	C4 Mismatched Year of Birth 0 = No 1 = Yes (CMS Data versus Survey Response)
198	C4BDGNDR	Num	8	1259	C4 Mismatched Gender 0 = No 1 = Yes (CMS Data versus Survey Response)
199	C4BDPRST	Num	8	1267	C4 Inconsistent-Female w/Prostate Cancer 0 = No 1 = Yes

VAR #	FIELD NAME	VAR TYPE	LENGTH/FORMAT	FIELD POSITION	LABEL DESCRIPTION
200	C4STNAME	Char	20	1275	C4 State Name from C4state SSA code
201	C4STABV	Char	2	1295	C4 State Abrv. from C4state SSA code
202	C4CTNAME	Char	25	1297	C4 County Name from C4county SSA code
203	C4RPTUNT	Char	6	1322	C4 Reporting Unit - Market Area Specific
204	C4RPT_CN	Char	6	1328	C4 Reporting Contract Number
205	C4RPT_MA	Char	1	1334	C4 Reporting Market Area N = North S = South E = East W = West O = Other
206	C4BASPCS	Num	8	1335	C4 Baseline PCS Score (MDE)
207	C4BASMCS	Num	8	1343	C4 Baseline MCS Score (MDE)
208	C4ADJPCS	Num	8	1351	C4 Adj. Physical Component Summary Score
209	C4ADJMCS	Num	8	1359	C4 Adj. Mental Component Summary Score
210	C4PCMNTU	Num	8	1367	C4 Mean PCS(Unadj) Total
211	C4MCMNTU	Num	8	1375	C4 Mean MCS(Unadj) Total
212	C4PCSDTU	Num	8	1383	C4 St Dev PCS(Unadj) Total
213	C4MCSDTU	Num	8	1391	C4 St Dev MCS(Unadj) Total
214	C4PCMNPU	Num	8	1399	C4 Mean PCS(Unadj) Plan
215	C4MCMNPU	Num	8	1407	C4 Mean MCS(Unadj) Plan
216	C4PCMXPUP	Num	8	1415	C4 Max PCS (Unadj) Plan
217	C4MCMXPUP	Num	8	1423	C4 Max MCS (Unadj) Plan
218	C4PCMIPUP	Num	8	1431	C4 Min PCS (Unadj) Plan
219	C4MCMIPUP	Num	8	1439	C4 Min MCS (Unadj) Plan
220	C4PCSDPUP	Num	8	1447	C4 St Dev PCS(Unadj) Plan
221	C4MCSDPUP	Num	8	1455	C4 St Dev MCS(Unadj) Plan
222	C4PCMNSU	Num	8	1463	C4 Mean PCS(Unadj) State
223	C4MCMNSU	Num	8	1471	C4 Mean MCS(Unadj) State
224	C4PCSDSU	Num	8	1479	C4 St Dev PCS(Unadj) State
225	C4MCSDSU	Num	8	1487	C4 St Dev MCS(Unadj) State
226	C4PCMNTA	Num	8	1495	C4 Mean PCS(Adj) Total
227	C4MCMNTA	Num	8	1503	C4 Mean MCS(Adj) Total
228	C4PCSDTA	Num	8	1511	C4 St Dev PCS(Adj) Total
229	C4MCSDTA	Num	8	1519	C4 St Dev MCS(Adj) Total
230	C4PCMNPAP	Num	8	1527	C4 Mean PCS(Adj) Plan
231	C4MCMNPAP	Num	8	1535	C4 Mean MCS(Adj) Plan
232	C4PCMXPAP	Num	8	1543	C4 Max PCS (Adj) Plan
233	C4MCMXPAP	Num	8	1551	C4 Max MCS (Adj) Plan
234	C4PCMIPAP	Num	8	1559	C4 Min PCS (Adj) Plan
235	C4MCMIPAP	Num	8	1567	C4 Min MCS (Adj) Plan
236	C4PCSDPAP	Num	8	1575	C4 St Dev PCS(Adj) Plan
237	C4MCSDPAP	Num	8	1583	C4 St Dev MCS(Adj) Plan
238	C4PCMNSAP	Num	8	1591	C4 Mean PCS(Adj) State
239	C4MCMNSAP	Num	8	1599	C4 Mean MCS(Adj) State
240	C4PCSDSAP	Num	8	1607	C4 St Dev PCS(Adj) State
241	C4MCSDSAP	Num	8	1615	C4 St Dev MCS(Adj) State

The following table is an alphabetic description of the file layout for the *Cohort IV Baseline* Electronic Data File.

VAR #	FIELD NAME	VAR TYPE	LENGTH/FORMAT	FIELD POSITION	LABEL DESCRIPTION
31	C4ACCRDT	Char	8	315	C4 Accretion Date
97	C4ACDING	Num	4	637	C4 Q19 Acid Indigestion
17	C4ADDRS1	Char	22	140	C4 Address Line 1
18	C4ADDRS2	Char	22	162	C4 Address Line 2
19	C4ADDRS3	Char	22	184	C4 Address Line 3
20	C4ADDRS4	Char	22	206	C4 Address Line 4
21	C4ADDRS5	Char	22	228	C4 Address Line 5
22	C4ADDRS6	Char	22	250	C4 Address Line 6
209	C4ADJMCS	Num	8	1359	C4 Adj. Mental Component Summary Score
208	C4ADJPCS	Num	8	1351	C4 Adj. Physical Component Summary Score
180	C4AGE	Num	8	1115	C4 Age (Continuous)
182	C4AGEGP2	Num	8	1131	C4 Age Groups for Baseline Report 1 = C4AGE ≥ 65 and <70 2 = C4AGE ≥ 70 and <75 3 = C4AGE ≥ 75 and <80 4 = C4AGE ≥ 80 (where C4AGE ≥ 65)
181	C4AGEGRP	Num	8	1123	C4 Age Groups 1 = Less than 55 2 = 55 to 64 3 = 65 to 74 4 = 75 or older
102	C4AMI	Num	4	657	C4 Q24 AMI
184	C4ANALYT	Num	8	1147	C4 Analytic Indicator 0 = No 1 = Yes (where C4AGE ≥ 65 and has PCS and/or MCS)
100	C4ANGCAD	Num	4	649	C4 Q22 Angina or CAD
89	C4ANKSWL	Num	4	605	C4 Q15b Ankles Swell
111	C4ANYCAN	Num	4	693	C4 Q33 Any Cancer
112	C4ARTHPN	Num	4	697	C4 Q34 Level arth pain
73	C4ASHLTH	Num	4	541	C4 Q11b As Healthy
108	C4ATHHAN	Num	4	681	C4 Q30 Arthritis hand or wrist
107	C4ATHHIP	Num	4	677	C4 Q29 Arthritis hip or knee
117	C4BACKPN	Num	4	717	C4 Q36 Back Pain Interferes
207	C4BASMCS	Num	8	1343	C4 Baseline MCS Score (MDE)
206	C4BASPCS	Num	8	1335	C4 Baseline PCS Score (MDE)
51	C4BATHDR	Num	4	453	C4 Q3j Bathing or dressing
197	C4BDBRTH	Num	8	1251	C4 Mismatched Year of Birth 0 = No 1 = Yes (CMS Data versus Survey Response)
195	C4BDCNUM	Num	8	1235	C4 Mismatched Contract Number 0 = No 1 = Yes (Plan ID versus Contract Number)

VAR #	FIELD NAME	VAR TYPE	LENGTH/FORMAT	FIELD POSITION	LABEL DESCRIPTION
198	C4BDGNDR	Num	8	1259	C4 Mismatched Gender 0 = No 1 = Yes (CMS Data versus Survey Response)
199	C4BDPRST	Num	8	1267	C4 Inconsistent-Female w/Prostate Cancer 0 = No 1 = Yes
196	C4BDRACE	Num	8	1243	C4 Mismatched Race 0 = No 1 = Yes (CMS Data versus Survey Response)
47	C4BEND	Num	4	437	C4 Q3f Bending
67	C4BLSAD	Num	4	517	C4 Q9f Blue/Sad
115	C4BRSTCA	Num	4	709	C4 Q35c Breast Cancer
127	C4BRTHYR	Char	4	757	C4 Q46 Survey-Birth Yr
101	C4CHF	Num	4	653	C4 Q23 CHF
46	C4CLMBON	Num	4	433	C4 Q3e Climb one flight
45	C4CLMBSV	Num	4	429	C4 Q3d Climb several flights
148	C4CMPFLG	Num	8	888	C4 Name Provided (Q56) flag
122	C4CMPHTH	Num	4	737	C4 Q41 Hth compared to peers
156	C4CMPSRV	Num	8	944	C4 Completed Survey (80% Criteria) 0 = No 1 = Yes
3	C4CNTRNM	Char	5	9	C4 Contract Number
113	C4COLNCA	Num	4	701	C4 Q35a Colon Cancer
186	C4COMO	Num	8	1163	C4 Number of Comorbids
187	C4COMOCT	Num	8	1171	C4 Comorbid Categories 1 = 0 or 1 Condition (s) 2 = 2 or 3 Conditions 3 = 4 or more Conditions
105	C4COPD_E	Num	4	669	C4 Q27 Emphy. Asthma COPD
24	C4COUNTY	Char	3	274	C4 County SSA Code
82	C4CPNEXR	Num	4	577	C4 Q13a Chest Pain-Exercise
83	C4CPNRST	Num	4	581	C4 Q13b Chest Pain-Resting
202	C4CTNAME	Char	25	1297	C4 County Name from C4county SSA code
98	C4CTRURN	Num	4	641	C4 Q20 Difclty cntrl urination
146	C4DBLDTY	Num	8	872	C4 Double Duty (in C4B and R2) 0 = No 1 = Yes
92	C4DECHEL	Num	4	617	C4 Q15e Sores that dont heal
91	C4DECSNS	Num	4	613	C4 Q15d Decreased feel hot/cold
121	C4DEP2YR	Num	4	733	C4 Q40 Two + years depressed
120	C4DEPMCH	Num	4	729	C4 Q39 Depressed in past yr
185	C4DEPRSS	Num	8	1155	C4 Positive Depression Screen (Yes to any Question 38, 39, or 40)
110	C4DIABET	Num	4	689	C4 Q32 Diabetes
76	C4DIFBTH	Num	4	553	C4 Q12a Difclty Bathing
79	C4DIFCHR	Num	4	565	C4 Q12d Difclty Getng out chair
77	C4DIFDRS	Num	4	557	C4 Q12b Difclty Dressing
78	C4DIFEAT	Num	4	561	C4 Q12c Difclty Eating
81	C4DIFTOL	Num	4	573	C4 Q12f Difclty Using Toilet
80	C4DIFWLK	Num	4	569	C4 Q12e Difclty Walking

VAR #	FIELD NAME	VAR TYPE	LENGTH/FORMAT	FIELD POSITION	LABEL DESCRIPTION
64	C4DNDMPS	Num	4	505	C4 Q9c Down in Dumps
29	C4DOB	Char	8	299	C4 Date of Birth
30	C4DOD	Char	8	307	C4 Date of Death
125	C4DRSQT	Num	4	749	C4 Q44 Time since Quit smoke
192	C4EACLMT	Num	8/ MMDDYY10	1211	C4 Elapsed Date-Accretion Limit-3/01/01
57	C4EACMPL	Num	4	477	C4 Q5b Emotnl-Accomplished less
190	C4EACRDT	Num	8/ MMDDYY10	1195	C4 Elapsed Date of Accretion into Plan
56	C4ECUTTM	Num	4	473	C4 Q5a Emotnl-cut down time
179	C4EDOB	Num	8/ MMDDYY10	1107	C4 Elapsed Date of Birth
188	C4EDOD	Num	8/ MMDDYY10	1179	C4 Elapsed Date of Death
132	C4EDUC	Num	4	777	C4 Q51 Education Level
66	C4ENERGY	Num	4	513	C4 Q9e Energy
194	C4ENRCAT	Num	8	1227	C4 Enrollment Duration Categories 1 = Less than 6 Months 2 = 6 to 12 Months 3 = 13 to 36 Months 4 = 37 or more Months
193	C4ENRDUR	Num	8	1219	C4 Enrollment Duration (Months)
58	C4ENTCRF	Num	4	481	C4 Q5c Emotnl-Not careful
33	C4ESRDID	Num	4	331	C4 ESRD Indicator
191	C4ESVDAT	Num	8/ MMDDYY10	1203	C4 Elapsed Date of Survey
189	C4ETRMDT	Num	8/ MMDDYY10	1187	C4 Elapsed Date of Termination from Plan
153	C4EVER	Num	8	920	C4 Indicator EVERCARE Plan
119	C4FELTSD	Num	4	725	C4 Q38 Sad/Blue in past 2 wks
14	C4FNAME	Char	15	112	C4 First Name
137	C4FNMCOMP	Char	15	797	C4 Q56 First Name Person comp
62	C4FULPEP	Num	4	497	C4 Q9a Pep/Life
28	C4GENDER	Num	4	295	C4 Gender (CMS)
40	C4GENHHTH	Num	4	409	C4 Q1 General Health Question
106	C4GI_ETC	Num	4	673	C4 Q28 GI Problems
69	C4HAPPY	Num	4	525	C4 Q9h Happy
140	C4HHINC	Num	4	847	C4 Q57 Household Income
99	C4HIGHBP	Num	4	645	C4 Q21 HTN or High BP
129	C4HISPAN	Num	4	765	C4 Q48 Hispanic Y/N
133	C4HMOWN	Num	4	781	C4 Q52 Homeowner Status
35	C4HOSPIC	Num	4	339	C4 Hospice Status
96	C4HRMOST	Num	4	633	C4 Q18 Hear most things
75	C4HTEXT	Num	4	549	C4 Q11d Health Excellent
9	C4HHTID	Char	7	71	C4 NCQA Healthcare Organization ID

VAR #	FIELD NAME	VAR TYPE	LENGTH/ FORMAT	FIELD POSITION	LABEL DESCRIPTION
41	C4HTRN	Num	4	413	C4 Q2 Health Transition Question
74	C4HTRWSE	Num	4	545	C4 Q11c Expt Hth to get worse
34	C4INSTUT	Num	4	335	C4 Institutional Status
157	C4INVALD	Num	8	952	C4 Ineligible by Plan
154	C4INVSrv	Num	8	928	C4 Invalid Survey
44	C4LIFT	Num	4	425	C4 Q3c Lift carry groceries
16	C4LNAME	Char	12	128	C4 Last Name
5	C4LNBUSS	Char	1	64	C4 Line of Business
139	C4LNMCOMP	Char	20	827	C4 Q56 Last Name Person comp
94	C4LSTTLK	Num	4	625	C4 Q16b Ever lost ability to talk
114	C4LUNGCA	Num	4	705	C4 Q35b Lung Cancer
160	C4MALNUM	Num	8	976	C4 Mail Completed Surveys by Plan
164	C4MALRR	Num	8	1008	C4 Mail Response Rate by Plan
131	C4MARITL	Num	4	773	C4 Q50 Marital Status
7	C4MARKET	Char	1	66	C4 Market Area
235	C4MCMIPA	Num	8	1567	C4 Min MCS (Adj) Plan
219	C4MCMIPU	Num	8	1439	C4 Min MCS (Unadj) Plan
231	C4MCMNPA	Num	8	1535	C4 Mean MCS(Adj) Plan
215	C4MCMNPU	Num	8	1407	C4 Mean MCS(Unadj) Plan
239	C4MCMNSA	Num	8	1599	C4 Mean MCS(Adj) State
223	C4MCMNSU	Num	8	1471	C4 Mean MCS(Unadj) State
227	C4MCMNTA	Num	8	1503	C4 Mean MCS(Adj) Total
211	C4MCMNTU	Num	8	1375	C4 Mean MCS(Unadj) Total
233	C4MCMXPA	Num	8	1551	C4 Max MCS (Adj) Plan
217	C4MCMXPU	Num	8	1423	C4 Max MCS (Unadj) Plan
150	C4MCONUM	Num	4	904	C4 MCO Provide Phone Num
237	C4MCS DPA	Num	8	1583	C4 St Dev MCS(Adj) Plan
221	C4MCS DPU	Num	8	1455	C4 St Dev MCS(Unadj) Plan
241	C4MCS DSA	Num	8	1615	C4 St Dev MCS(Adj) State
225	C4MCS DSU	Num	8	1487	C4 St Dev MCS(Unadj) State
229	C4MCS DTA	Num	8	1519	C4 St Dev MCS(Adj) Total
213	C4MCS DTU	Num	8	1391	C4 St Dev MCS(Unadj) Total
135	C4MDSVPV	Num	4	789	C4 Q54 Medical Svcs Provided
36	C4MEDICD	Num	4	343	C4 Medicaid Status
15	C4MI	Char	1	127	C4 Middle Initial
138	C4MNMCOMP	Char	15	812	C4 Q56 Middle Name Person comp
43	C4MODACT	Num	4	421	C4 Q3b Moderate Activities
6	C4MODEL	Char	1	65	C4 Model Type
166	C4MONRPT	Char	8	1024	C4 Monthly Report - source CMS 06/01
63	C4NERVS	Num	4	501	C4 Q9b Nervous
88	C4NMBFET	Num	4	601	C4 Q15a Numbness Feet
118	C4NUMBLG	Num	4	721	C4 Q37 Numbness in leg
103	C4OTHRT	Num	4	661	C4 Q25 Other Heart cond.

VAR #	FIELD NAME	VAR TYPE	LENGTH/ FORMAT	FIELD POSITION	LABEL DESCRIPTION
152	C4PACE	Num	8	912	C4 Indicator PACE Plan
53	C4PACMPL	Num	4	461	C4 Q4b Phys-Accomplish less
93	C4PARLYS	Num	4	621	C4 Q16a Ever had paralysis
177	C4PATID	Num	8	1091	C4 Anonymous Patient ID 1 to 190,523
65	C4PCEFUL	Num	4	509	C4 Q9d Peaceful
234	C4PCMIPA	Num	8	1559	C4 Min PCS (Adj) Plan
218	C4PCMIPU	Num	8	1431	C4 Min PCS (Unadj) Plan
230	C4PCMNPA	Num	8	1527	C4 Mean PCS(Adj) Plan
214	C4PCMNPU	Num	8	1399	C4 Mean PCS(Unadj) Plan
238	C4PCMNSA	Num	8	1591	C4 Mean PCS(Adj) State
222	C4PCMNSU	Num	8	1463	C4 Mean PCS(Unadj) State
226	C4PCMNTA	Num	8	1495	C4 Mean PCS(Adj) Total
210	C4PCMNTU	Num	8	1367	C4 Mean PCS(Unadj) Total
232	C4PCMXP	Num	8	1543	C4 Max PCS (Adj) Plan
216	C4PCMXP	Num	8	1415	C4 Max PCS (Unadj) Plan
236	C4PCSDPA	Num	8	1575	C4 St Dev PCS(Adj) Plan
220	C4PCSDPU	Num	8	1447	C4 St Dev PCS(Unadj) Plan
240	C4PCSDSA	Num	8	1607	C4 St Dev PCS(Adj) State
224	C4PCSDSU	Num	8	1479	C4 St Dev PCS(Unadj) State
228	C4PCSDTA	Num	8	1511	C4 St Dev PCS(Adj) Total
212	C4PCSDTU	Num	8	1383	C4 St Dev PCS(Unadj) Total
155	C4PCTCMP	Num	8	936	C4 Percent of Survey Completed
52	C4PCUTTM	Num	4	457	C4 Q4a Phys-Cut down time
55	C4PDIFWK	Num	4	469	C4 Q4d Difficulty w/work
178	C4PLAN	Num	8	1099	C4 Anonymous Plan ID 1 to 197
13	C4PLANID	Char	6	106	C4 Plan ID
4	C4PLANNM	Char	50	14	C4 Plan Name
171	C4PLANST	Char	2	1053	C4 Plan State - source CMS 06/01
170	C4PLDESC	Char	5	1048	C4 Plan Description - source CMS 06/01
175	C4PLDUR	Num	8	1075	C4 Duration of Plan Contract - Years
168	C4PLMODL	Char	5	1035	C4 Plan Model - source CMS 06/01
54	C4PLMTKW	Num	4	465	C4 Q4c Lmtd in Kind of work
176	C4PLNDCT	Num	8	1083	C4 Duration of Plan Contract Categories 1 = Less than 1 Year 2 = 1 to 4.9 Years 3 = 5.0 to 9.9 Years 4 = 10 or more Years
169	C4PLPOP	Char	8	1040	C4 Plan Population - source CMS 06/01
173	C4PLREG	Num	8	1059	C4 Plan CMS Region - source CMS 06/01
174	C4PLSTDT	Num	8/ MMDDYY10	1067	C4 Contract Start Date -source CMS 06/01
172	C4PLTXST	Char	4	1055	C4 Plan Tax Status - source CMS 06/01
167	C4PLTYPE	Char	3	1032	C4 Plan Type - source CMS 06/01
167	C4PLTYPE	Char	3	1032	C4 Plan Type - source CMS 06/01

VAR #	FIELD NAME	VAR TYPE	LENGTH/FORMAT	FIELD POSITION	LABEL DESCRIPTION
61	C4PNINTF	Num	4	493	C4 Q8 Pain-Interfere Work
60	C4PNMAGT	Num	4	489	C4 Q7 Pain-Magnitude
116	C4PROSCA	Num	4	713	C4 Q35d Prostate Cancer
38	C4PROTIF	Num	8	351	C4 Protocol Identifier Flag
145	C4PROXST	Num	4	868	C4 Proxy Status
126	C4QSMKAD	Num	4	753	C4 Q45 Freq antismoke advice
27	C4RACE	Num	4	291	C4 Race (CMS)
183	C4RACEGP	Num	8	1139	C4 Race Groups for Baseline Report 1 = White 2 = Black 3 = Other (CMS data used – C4Race)
95	C4RDNEWSP	Num	4	629	C4 Q17 See to read newspaper
1	C4RECID	Char	1	0	C4 Record ID
39	C4RESPEC2	Char	50	359	C4 Respondent to Cohort 2 Survey
142	C4RNDNUM	Char	2	854	C4 Round Survey Obtained
147	C4ROUND	Num	8	880	C4 Round Data Submitted (1 or 2)
204	C4RPT_CN	Char	6	1328	C4 Reporting Contract Number
205	C4RPT_MA	Char	1	1334	C4 Reporting Market Area
203	C4RPTUNT	Char	6	1322	C4 Reporting Market Area N = North S = South E = East W = West O = Other
2	C4RPTYR	Num	8	1	C4 Report Year
37	C4RSENT	Num	4	347	C4 Reason for Entitlement
162	C4RSPDEN	Num	8	992	C4 Sample minus invalids by Plan
159	C4RSPNUM	Num	8	968	C4 Completed Surveys by Plan
163	C4RSPON	Num	8	1000	C4 Response Rate by Plan
134	C4RTRCOM	Num	4	785	C4 Q53 Retirement Community
158	C4SAMPLE	Num	8	960	C4 Sample Size by Plan
109	C4SCIATC	Num	4	685	C4 Q31 Sciatica
72	C4SCKESY	Num	4	537	C4 Q11a Sick Easier
71	C4SCLACT	Num	4	533	C4 Q10 Social-time
123	C4SMK100	Num	4	741	C4 Q42 Ever smoke 100 cigs
124	C4SMKFRQ	Num	4	745	C4 Q43 Curnt smoke freq
84	C4SOBFLT	Num	4	585	C4 Q14a Short Brth lying flat
85	C4SOBSIT	Num	4	589	C4 Q14b Short Brth sittng/restng
87	C4SOBSTR	Num	4	597	C4 Q14d Short Brth Climb 1 flight stairs
86	C4SOBWLK	Num	4	593	C4 Q14c Short Brth Walk<1 block
59	C4SOCLMT	Num	4	485	C4 Q6 Social-extent
151	C4SPANFL	Num	4	908	C4 Spanish Protocol Flag
11	C4SPECID	Char	10	84	C4 NCQA Special Area ID
141	C4SRVDSP	Char	3	851	C4 Survey Disposition
26	C4SSN	Char	9	282	C4 Social Security Number
201	C4STABV	Char	2	1295	C4 State Abrv. from C4state SSA code
23	C4STATE	Char	2	272	C4 State SSA Code
200	C4STNAME	Char	20	1275	C4 State Name from C4state SSA code

VAR #	FIELD NAME	VAR TYPE	LENGTH/FORMAT	FIELD POSITION	LABEL DESCRIPTION
104	C4STROKE	Num	4	665	C4 Q26 Stroke
10	C4SUBID	Char	6	78	C4 NCQA Submission ID
128	C4SV_GND	Num	4	761	C4 Q47 Survey-Gender
130	C4SV_RAC	Num	4	769	C4 Q49 Survey-Race
144	C4SVDATE	Char	8	860	C4 Date Survey Completed
143	C4SVLANG	Num	4	856	C4 Survey Language 1 = English 2 = Spanish 3 = Not Applicable
161	C4TELNUM	Num	8	984	C4 Tele Completed Surveys by Plan
165	C4TELRR	Num	8	1016	C4 Tele Response Rate by Plan
32	C4TERMDT	Char	8	323	C4 Termination Date
90	C4TINGFT	Num	4	609	C4 Q15c Tingling Burning feet
70	C4TIRED	Num	4	529	C4 Q9i Tired
8	C4VENDOR	Num	4	67	C4 Vendor 1 = DSS 2 = GHS 3 = Solucient 4 = Market Facts
42	C4VIGACT	Num	4	417	C4 Q3a Vigorous Activities
149	C4VUCATI	Char	8	896	C4 Vendor Unique CATI ID
136	C4WHOCMP	Num	4	793	C4 Q55 Who completed Survey
50	C4WLK1BK	Num	4	449	C4 Q3i Walk 1 block
49	C4WLKBKS	Num	4	445	C4 Q3h Walk several blocks
48	C4WLKMI	Num	4	441	C4 Q3g Walk > 1 mile
68	C4WRNOUT	Num	4	521	C4 Q9g Worn Out
25	C4ZIPCOD	Char	5	277	C4 Zip Code
12	HICNUM	Char	12	94	HIC Number